UDC 534.86:
621.376.56

1969
THE QUEEN'S AWARD
TO INDUSTRY

RESEARCH DEPARTMENT REPORT

Pulse code modulation for high-quality signal distribution: quantizing distortion at very low signal levels

No. 1970/18

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

RESEARCH DEPARTMENT

PULSE CODE MODULATION FOR HIGH QUALITY SOUND DISTRIBUTION: QUANTIZING DISTORTION AT VERY LOW SIGNAL LEVELS

Research Department Report No. **1970/18**
UDC 534.86
621.376.56

It uses SI units in accordance with B.S. document PD 5686.

Work covered by this report was undertaken by the BBC Research Department for the BBC and the ITA

M.G. Croll, B.Sc., A.R.C.S.

D. Maurice
Head of Research Department

(EL-40)

Research Department Report No. 1970/18

PULSE CODE MODULATION FOR HIGH-QUALITY SOUND DISTRIBUTION: QUANTIZING DISTORTION AT VERY LOW SIGNAL LEVELS

(EL-40)

June 1970

Research Department Report No. 1970/18
UDC 534.86
621.376.56

PULSE CODE MODULATION FOR HIGH QUALITY SOUND DISTRIBUTION: QUANTIZING DISTORTION AT VERY LOW SIGNAL LEVELS

SUMMARY

In p.c.m. systems, distortion of very low level signals is known to occur, and in high-quality programme distribution it is the most objectionable signal degradation caused by this kind of system.

This report describes a series of subjective tests which show that it is necessary to use 14 bits per sample to code the audio signal in order to make this form of distortion, often called 'granular distortion', inaudible. However, granular distortion can be substantially eliminated by the addition of small agitating signals at the input of the system with only a slight decrease in the overall signal-to-noise ratio. It is recommended that this artifice be employed to reduce the number of bits per sample to 13.

1. INTRODUCTION

1.1. General

In a p.c.m. system the analogue input is quantized into a finite number of discrete amplitudes. Through this quantizing process the instantaneous output signal differs from the input signal by arbitrary amounts up to half a quantizing step. Hence the fewer quantizing levels explored by the input signal the greater the percentage distortion of the reconstituted output signal.

When the input signal explores sufficient quantizing levels the distortion sounds like random noise added to the signal and is referred to in this report as 'quantizing hiss.' In pauses in programme, when the residual input signal is less than one quantizing step — a situation known as the idling condition — the output contains most noise when the mean input voltage lies at a decision level of the analogue-to-digital converter (a.d.c.) and least when it lies midway between two such levels. This form of noise will be referred to as 'idling noise.' When the input signal is larger than one quantizing step, but not of sufficient level to produce a steady quantizing hiss, the resultant effect is referred to as 'granular distortion.'[1]

1.2. Granular Distortion

In this report the term 'granular distortion' is used to describe the wide range of different audible distortions, occurring with low-level signals, which are not classifiable either as quantizing hiss or idling noise.

If the input signal is nearly large enough to produce a steady quantizing hiss, the hiss is modulated by the programme. Smaller signals are audibly distorted and contain gaps as the quietest passages of the programme are lost between adjacent quantizing levels and the idling condition is approached.

The extremely low programme levels which produce granular distortion at the output of a binary p.c.m. system having more than 2048 quantizing levels, corresponding to 11 bits, commonly occur towards the end of a slow fade or when music is used as a low-level background sound effect.

As the percentage distortion of the transmitted signal is an inverse function of the number of quantizing levels explored, the distortion can be reduced if more bits — and hence more quantizing levels — are provided. The tests described in this report set out to determine the number of bits that must be used to describe the analogue signal in order to make granular distortions inaudible and to try some artifices which might reduce this requirement.

2. A TEST TO DETERMINE THE NUMBER OF BITS NECESSARY TO MAKE GRANULAR DISTORTION INAUDIBLE

2.1. Choice of Test Material

The p.c.m. system used for the tests comprised a 10-bit ramp a.d.c. and a 10-bit current-adding digital-to-analogue converter (d.a.c.), both operating at a 30 kHz sampling rate. Using this apparatus tests

were made to discover the type of programme material most affected by granular distortion. With speech, orchestral or violin music, the ill effects were masked to some extent by the programme itself. Granular distortion was most noticeable on staccato piano music, and a tape recording of some piano duets was therefore chosen for the tests. This recording had a very limited dynamic range and was therefore suitable for continuous tests.

2.2. The Simulation of a p.c.m. System having a Large Number of Bits

The objective of this test was to establish the minimum number of bits per sample which would provide a system free of impairment by granular distortion. This form of distortion is produced by low level programme and piano music is more likely to be impaired than most other types of programme material. It was therefore decided that if a given number of bits per sample would permit piano music to be slowly faded from a low level to inaudibility without inducing perceptible granular distortion then that number of bits would be satisfactory for high quality music distribution.

A fader at the audio input to the p.c.m. system was therefore provided to enable the observer to explore the range of low signal levels in which granular distortion might be heard.

In order to simplify the experiment, a 10-bit p.c.m. system was used and advantage was taken of the fact that an attenuator in the audio output of the p.c.m. system enables systems using larger numbers of bits per sample to be simulated, provided that, as in this case, only low level programme is required. For example:— a given listening level results from an input signal level which excurses over 100 quantizing levels; if the input level is increased and the output is attenuated, both by 6 dB, the listening level will be unchanged but the signal will have occupied 200 quantum levels. Thus, by placing an attenuator at the output of the p.c.m. system, the effect of systems having greater numbers of bits per sample can be simulated at the rate of one extra bit for each 6 dB of attenuation. By this means the 10-bit system was able to simulate systems up to 15 bits per sample, provided the level of signal was sufficiently low.

The arrangement of equipment is shown in Fig. 1.

Attenuator 1 was used to provide and fade very low-level programme, attenuator 2 to simulate greater numbers of digits and attenuator 3 to set the listening level. All the attenuators were variable in steps of 1 dB.

2.3. The Subjective Experiment

An excerpt of piano music was played and the observer was asked to set his own preferred listening level by adjusting attenuator 3, attenuators 1 and 2 being set at zero attenuation. He was then instructed to reduce the input to the p.c.m. system slowly to inaudibility by adjusting attenuator 1 while listening for distortion, and to repeat this process with various settings of attenuator 2. The setting of attenuator 2 at which the distortions were just inaudible to the observer was noted. The condition 'just inaudible' was defined by the least attenuation in attenuator 2 at which the distortions were not audible.

The peak listening levels in the test were measured by substituting for the programme source an octave band of white noise centred on 1 kHz. The noise level was set to peak to +8 dB on the P.P.M. with attenuators 1 and 2 set at zero loss. The sound output level from the loudspeaker was measured at the listening position for the settings of attenuator 3 chosen by the observers, using a sound level meter and an octave band-pass filter (to BS.2475 : 1964) centred on 1 kHz. The peak listening levels ranged from 80—90 dB with respect to 0·00002 Newtons/metre^2.

2.4. Results

Sixteen observers took part in this test. The mean attenuation that was set by the observers on attenuator 2 was: 21·9 dB (S.D. 3·7 dB).* This figure is equivalent to the addition of a further $3^2/3$ bits to the 10-bit p.c.m. system, calculated on the basis that for every 6 dB attenuation inserted an extra bit is simulated.

The results of this test are also shown plotted on arithmetical probability paper in Fig. 2, from which it can be seen that 80% of the observers would not hear the granular distortion of a system using 14 binary bits, and 50% of the observers would not hear the granular distortion of a 13½ bit system.

* Note: In this report results are quoted x (S D.y) where x = mean value; y = standard deviation of the set of results.

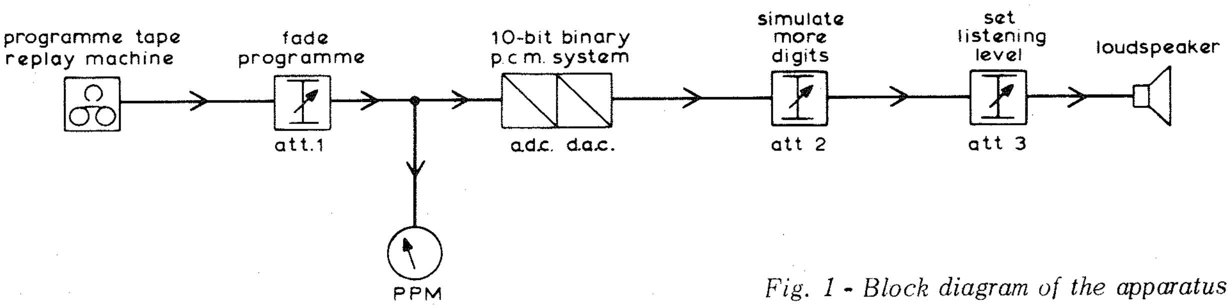

Fig. 1 - Block diagram of the apparatus

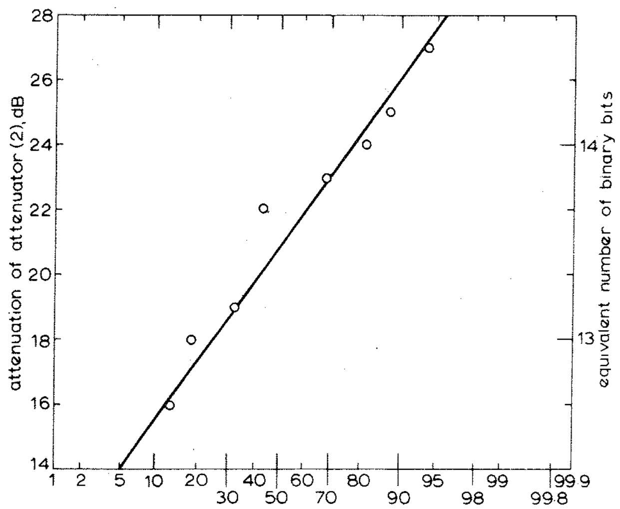

Fig. 2 - The number of bits necessary to make granular distortion inaudible

3. THE USE OF ARTIFICES TO REDUCE GRANULAR DISTORTION

The subjective tests described in the previous section were based on the use of a simple p.c.m. system without companding, pre-emphasis or other means of alleviating the impairments resulting from quantization of the audio signal.

Earlier work[2,3] has shown that instantaneous companding, common in p.c.m. telephony, is not able to provide a worthwhile improvement in systems for high-quality music transmission without introducing other and equally disturbing forms of distortion.

Compandors with slow response times, sometimes called syllabic compandors, are more acceptable for music transmission systems and a specially developed form using a pilot tone has been used effectively with the Sound-in-Syncs system of distributing television sound signals.[4] Although these compandors are acceptable for monophonic transmission they exhibit imperfections when used with audio signals for stereophonic programmes. There are available more sophisticated compandors which can meet the stringent requirements of stereophonic signals. They can make improvements of the order of 12 dB (equivalent to a saving of 2 bits/sample) but their cost is high. Thus the cost and complexity of a syllabic compandor is only justified when there are very strong reasons for using the lowest possible number of bits/sample.

However the use of pre-emphasis is neither costly nor prone to cause distortion and the effectiveness of this technique was tested experimentally.

The use of added signals, either random or deterministic has been successfully applied to reduce the number of bits/sample of p.c.m. systems for video signals,[5] and these techniques were also tested in the experiments reported below; a 10-bit p.c.m. system was used so that the granular distortion would be clearly audible to all the observers.

3.1. An Attempt to Reduce Granular Distortion by Pre- and De-emphasis

Pre- and de-emphasis has previously been found effective in reducing the output noise of analogue sound systems and these networks were therefore tried as a means of reducing the granular distortion of a p.c.m. system. A measurement of the reduction of granular distortion due to pre- and de-emphasis was made by reducing the effective number of digits through suitable adjustment of attenuators before and after the system, and comparing the output subjectively with that from the system without pre- and de-emphasis. Hence the reduction of distortion was expressed in terms of dB additional attenuation inserted before the system.

(a) The use of 50 μsec pre-emphasis characteristics

The pre- and de-emphasis networks were, for this test, set to have 0 dB gain at 100 Hz.

Some reduction in the effect of granular distortion was observed and was judged to be about 3 to 4 dB for piano music. For a programme which consists mainly of high-frequency components, i.e. glockenspiel, the improvement was 4 to 6 dB.

It has been found that[6] in order to use 50 μsec networks in a broadcast transmission system, the pre-emphasis network must be set to have approximately 4 dB attenuation at low frequencies to allow for the increase in level at high frequencies. Therefore, little useful reduction in granular distortion could be expected if 50 μsec networks were used.

(b) The C.C.I.T.T. Characteristics

The C.C.I.T.T. pre- and de-emphasis characteristics have a stepped frequency response, the gain rising and falling respectively 18 dB between 100 Hz and 11 kHz; in this test the pre- and de-emphasis networks were each set to have 0 dB gain at 100 Hz.

Using these networks there was a considerable reduction in granular distortion. The improvement was judged to be 16 dB although gaps in the signal were perceptible at very low signal levels. When these gaps were included in the assessment of the impairment then the improvement was judged to be only 11 dB.

However, it has been found that the overall increase in signal level due to pre-emphasis would require that the pre-emphasis network be set to have at least 10 dB attenuation at low frequencies to avoid applying an excessive signal level to the input of the p.c.m. system. Therefore, the overall advantage would not be more than 1 dB.

3.2. An Attempt to Mask Granular Distortion by Adding White Noise to the Output

In the test referred to in Section 3 it was found that 14 bits per sample would be necessary to make granular distortions inaudible to most observers; the resultant signal-to-quantizing hiss ratio would be 77 dB.* For a high-quality audio-frequency distribution system, a figure of 63 dB** is considered adequate. If, therefore, quantizing hiss and not granular distortion were the main consideration then a system equivalent to $11^2/3$ digits only need be used; 13 digits, giving a signal-to-noise ratio of 71 dB, would then enable four such systems to be operated in tandem with 2 dB in hand. Hence, if masking of granular distortion could be effected even at a slight reduction of the signal-to-noise ratio then only 13 bits per word need be used to code the audio signal.

An attempt was made to mask granular distortion by the addition of white noise to the output of the system. It was found necessary to add white noise at a level 4 dB higher than the quantizing hiss of the system before a complete masking of granular distortion was effected. The resultant signal-to-noise ratio of the total system was degraded by about 5 dB and this solution was therefore not further considered.

3.3. Reduction of Granular Distortion Through the Addition of 'dither' Waveforms to the Input of the p.c.m. System

The granular distortion of quantized audio can be likened to the appearance of spurious contours when video signals are coarsely quantized; both are the resultant subjective effects of insufficient quantizing levels being explored by the signal. The visibility of such contours on a TV picture can be effectively reduced by adding to the signal before quantizing a small-amplitude 'dither' waveform.[5]

When a dither waveform is added to the normal input of a p.c.m. system the total input signal excurses over more quantizing levels within the p.c.m. system. Hence the distortion of the system may be reduced, although the dither waveform may increase the noise output of the system.

The dither waveforms used here were half-sampling frequency square waves and/or white noise and, unlike the noise in Section 3.2. were added to the audio signal at the input to the a.d.c.

(a) Half-sampling Frequency Square Waves

Provision was made within the analogue-to-digital converter for half-sampling-frequency square waves to be generated and added to the input analogue signal. The peak-to-peak amplitude of the square waves was adjusted to equal half the quantizing step of the coder, so that for alternate samples the quantizing levels were in effect midway between the quantizing levels used for the previous sample.

When this square wave was present the nature of the granular distortion was changed in that its impairment of the transmitted signal was reduced, but additional high-frequency components were produced. The nett subjective improvement was less than 2 dB.

(b) White Noise

White noise was added at the input of the coder and it was found that the granular distortion of the system could be rendered inaudible. A subjective test was carried out in which 12 observers were asked to add just sufficient noise to make the granular distortion of the 10-bit p.c.m. system inaudible. The noise added was measured relative to the quantizing hiss of the system at higher signal levels.

The results shown in Fig. 3(*a*) indicate that granular distortion can be rendered inaudible to 50% of the observers if white noise, 2 dB higher in level than the quantizing hiss, is added to the signal input to the p.c.m. system.

(c) Half-sampling-frequency Square Waves Plus White Noise

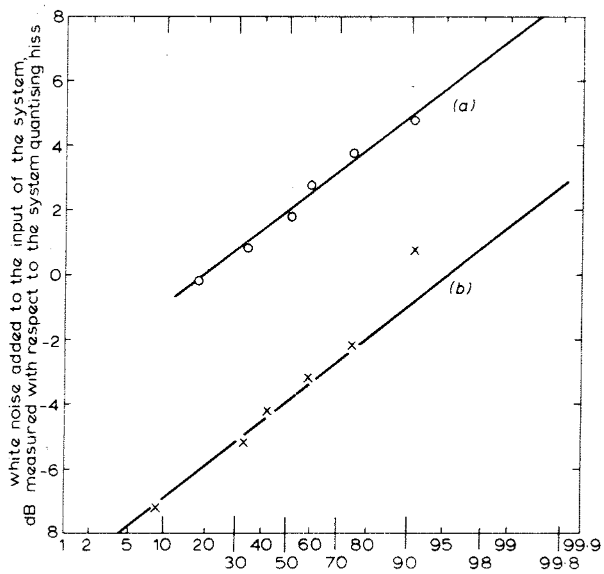

Fig. 3 - The quantity of white noise that must be added at the input to mask granular distortion

(*a*) White noise added alone
(*b*) White noise and half sampling frequency square waves added

* Peak signal to peak weighted noise

** The signal-to-noise ratio for the complete broadcasting chain should ideally be at least 60 dB;[7] assuming that up to half the total noise power is contributed by the studio equipment and the transmitter/receiver combination, the signal-to-noise ratio of the distribution links alone should not be less than 63 dB

The amount of white noise that must be added at the input of the p.c.m. system to render inaudible the granular distortion of the system was measured when the square waves were also added at the input. The subjective test was conducted similarly to that described in 3.3.(b).

The results, shown in Fig. 3(*b*), indicate that, in the presence of half-sampling-frequency square waves, white noise can render granular distortion inaudible to 50% of the observers if it is at a level 4 dB lower than the quantizing hiss of the system. Hence the action of the square waves has been to reduce by 6 dB the applied noise level necessary to mask granular distortion. If both artifices were simultaneously applied to a 13-bit p.c.m. system the granular distortion of the system would be at a level 18 dB lower than that presented to the observers in this test and could be regarded as inaudible. The signal-to-noise ratio of the 13-bit system would be degraded by only 2 dB. In special cases where the signal-to-noise ratio might be required to be even better, standard noise reduction techniques such as pre- and de-emphasis could be applied. (For noise reduction in p.c.m. these techniques are as effective as in analogue systems).

4. CONCLUSIONS

(i) In order to render the granular distortion of a simple binary p.c.m. system substantially inaudible, 14-bits per sample must be used to code the signal, although from quantizing hiss considerations alone a 12-bit system is more than adequate for a single codec and a 13-bit system would enable 4 codecs to be operated in tandem.

(ii) Granular distortion is slightly reduced if half-sampling frequency square waves, with peak-to-peak amplitudes equal to half a quantizing step, are added to the input to the system. The distortion can be substantially eliminated if sufficient white noise is added at the input to the system; the amount of noise required is 6 dB less if the square waves are also added. If both artifices are used the signal-to-noise ratio of the system need only be impaired by some 2 dB.

(iii) Other means of reducing the number of bits/sample are less attractive; pre-emphasis offering, at best, a reduction of only 2 dB or $^1/_3$ of a bit and compandors offering larger savings of up to 12 dB or 2-bits but at considerable cost.

The use of compandors would only be justified if the advantages of saving 2-bits/sample were considerably greater than the cost of providing the companding equipment.

It is therefore recommended that the minimum number of bits-per-sample in a p.c.m. system for use by high quality sound signals should be 13. The addition of half-sampling frequency square waves together with noise will ensure freedom from distortion and the noise level of the system will allow four coding and decoding processes to be put in tandem without exceeding the recommended figure of 63 dB.

5. REFERENCES

1. RICHARDS, D.L. 1968. Transmission performance of telephone networks containing p.c.m. links. *Proc. Instn elect. Engrs,* 1968, **115**, 9, pp. 1245–1258.

2. Pulse-code modulation for high-quality sound signal distribution: instantaneous compandors using diodes as non-linear elements. BBC Research Department Report No. EL-15, Serial No. 1968/7.

3. Pulse-code modulation for high-quality sound-signal distribution: a simulation of instantaneous companding. BBC Research Department Report No. 1969/26.

4. A frequency-dependent compandor system for high-quality sound signal distribution. BBC Research Department Report No. EL-21, Serial No. 1968/52.

5. LIMB, J.O. 1969. Design of dither waveforms for quantized visual signals. *Bell Syst. tech. J.,* 1969, **48**, 7, pp. 2555–2582.

6. HAYES, E.W. *and* PAGE, H. 1956. The B.B.C. sound broadcasting service on very high frequencies. *Proc. Instn elect. Engrs,* 1957, **104B**, 15, pp. 213–224.

7. The assessment of noise in audio frequency circuits. BBC Research Department Report No. EL-17, Serial No. 1968/8.

JMP/JUC